POSTER

What Can You Do For Me? The Discoverability of Intelligent Assistant Skills

MANVEER KALIRAI, University of Toronto, Toronto, ON, Canada

ANASTASIA KUZMINYKH, University of Toronto, Toronto, ON, Canada

Published: 24 April 2023

Citation in BibTeX format

UbiComp/ISWC '22: The 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing
September 11 - 15, 2022
Cambridge, United Kingdom

Conference Sponsors:
SIGMOBILE
SIGCHI

https://doi.org/10.1145/3544793.3560362
ISBN: 9781450394239

What Can You Do For Me? The Discoverability of Intelligent Assistant Skills

Manveer Kalirai
University of Toronto
Toronto, Canada

Anastasia Kuzminykh
University of Toronto
Toronto, Canada

ABSTRACT

A wide discrepancy exists between the range of available Intelligent Personal Assistant (IPA) skills and the range of skills users engage with regularly. One reason for this is the issue of low skill discoverability. From the literature, we understand that there are factors that can potentially enhance discoverability, such as context of use. The literature also signals that current discoverability strategies, which leverage such factors, are being challenged by users' privacy concerns, and rapid skill growth. Mindful of these challenges, we explore the ways users naturally group IPA skills as a possible springboard for discoverability. Our preliminary findings suggest that grouping skills based on users' functionality needs in different areas of life may be a viable direction toward advancing discoverability.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design**; **HCI theory, concepts and models**.

KEYWORDS

conversational agents, voice interface, skills, discoverability

ACM Reference Format:
Manveer Kalirai and Anastasia Kuzminykh. 2022. What Can You Do For Me? The Discoverability of Intelligent Assistant Skills. In *Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp/ISWC '22 Adjunct), September 11–15, 2022, Cambridge, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3544793.3560362

1 INTRODUCTION AND BACKGROUND

The adoption of smart speakers has surged in recent years, bounding from 76 million to 94 million owners in the United States between 2020 to 2021 [10]. The 'smartness' of smart speakers is ascribed to their Intelligent Personal Assistants (IPAs), the artificial agents with whom users interact whilst using the devices. IPAs are software-based entities that leverage artificial intelligence to assist users with a breadth of everyday tasks, such as setting timers, playing music, and answering general information queries.

The assistance provided by IPAs is channeled through an assemblage of functionalities for these systems called 'skills'. Alongside the growing adoption of smart speakers, the corpus of IPA skills is also growing. In 2021, there were reportedly over 80,000 skills available for popular IPA Alexa (Amazon); a considerable leap from 24,499 in 2017 [14]. However, despite the wide range of skills currently available to users, the range of skills with which users are actually engaging is far more narrow—users tend to engage with 3-10 skills on average [1]. This disparity can, in part, be explained as the result of low skill discoverability [6, 14], i.e. a user's ability to understand what functionalities, or skills, are available within their IPA system.

Through a systematic review of related research, we first identified three levels of factors that make the discoverability of IPA skills particularly *challenging*: modality-level factors, focused on the conversational nature of an IPA's interface [9][12]; skill-level factors, grouped around the inconsistent representation of skills [11][13]; and skill command-level factors, reflecting unnatural query formulations and cognitively-demanding command remembrance [5][8].

We then analyzed and grouped *discoverability-enhancing* factors, explored in previous works. These include: accounting for context of skill-use, which refers to holding relevance to the user task or activity being supported at the time of IPA-use [14]; level of user expertise, referring to the user journey in learning to use an IPA, and gaining familiarity with its available functionalities [3]; and user preferences for IPA proactiveness, i.e. for the proactive or reactive nature of skill recommendations [6]. Furthermore, these groups of factors seem to be interconnected, e.g. the desired level of IPA proactiveness when providing recommendations has been observed to vary based on: (i) users' learning progression [2], and (ii) users' privacy preferences [7].

Finally, as our understanding of the factors challenging or enhancing the discoverability of IPA skills continues to expand, work capitalizing on them continues to grow in tandem. For instance, recent *approaches to addressing skill discovery*, focus on providing tailored prompts that are context-specific as well as user-specific, being inferred from a range of real-time contextual and personal signals [4]. However, we suggest that, although promising, these current approaches in the skill discovery space are challenged by two aspects: (i) users' privacy concerns toward IPAs [7], and (ii) the rapid growth of the corpus of IPA skills [14].

To summarize, previous research suggests that to address discoverability challenges for conversational IPAs, e.g. due to the inconsistencies in the naming of skills and their functionalities, inconsistent activity scope of each skill, and low intuitiveness of required queries, the representation of skills should be meaningfully situated within the user's current activity, i.e. the task context. The understanding of what such context constitutes remains underdeveloped, yet, critical for enabling context-dependent skills presentation without continuous real-time data gathering. Correspondingly, in this project, we explore users' intuitive groupings of IPA skills as a mechanism for context-enhanced discoverability.

UbiComp/ISWC '22 Adjunct, September 11–15, 2022, Cambridge, United Kingdom

ACM ISBN 978-1-4503-9423-9/22/09.
https://doi.org/10.1145/3544793.3560362

2 INITIAL RESEARCH

At the first stage of this project, we analyze users' functionality needs and corresponding motivations for IPA-use as a basis for developing skill groupings. We focus particularly on collecting data during COVID-19 pandemic lockdowns in 2020 and 2021, as it marked the sudden transition of various areas of life into a single space—the home. Building on past research, we hypothesized that functionality needs and motivations for IPA-use and contexts of IPA-use would differ by area of life, and these unprecedented times presented a unique opportunity to analyze these differences.

To gain initial insight on emerging motivations and contexts of IPA-use during COVID-19, we developed a dataset of relevant online discussions, gathered through netnography. Our dataset was composed of publicly-accessible discussions that had been drawn from various online social media and news media platforms, including Reddit, Twitter, YouTube, CNET, etc. Our inclusion criteria required data points to reference use of smart speakers during the pandemic, and to have been published in the year 2020 or 2021. The resulting dataset included 189 data points (quotes) belonging to 166 conversations (maximum: 7 quotes per conversation; minimum: 1 quote per conversation), with a mean of 1.1 quotes per conversation.

All 189 quotes in the dataset were then thematically analyzed around motivations and contexts of IPA-use. The initial round of coding revealed the presence of five areas of use, these being: Parenting, Work, Housekeeping, Leisure, and Health. Subsequent rounds of coding were focused on surfacing motivations and contexts of IPA-use as well as specific IPA features discussed within these five themes. Ultimately, a total of 28 codes were identified across 12 contexts of use, and across five areas of life.

Our findings on the motivations and contexts for IPA-use during COVID-19 stay-at-home orders demonstrate varying functionality needs in five different areas of use. These areas of use were identified through a thematic analysis of our netnographic data, and included: **Parenting** (n=21), **Work** (n=22), **Housekeeping** (n=58), **Leisure** (n=72), and **Health** (n=18). Additionally, 12 subthemes around 'contexts of use' were identified across the above five areas of life. For **Parenting**, these contexts of use included: encouraging desired behaviours [e.g. Q146: "*Google Home does this to help kids wash their hands for the correct amount of time in the pandemic.*"], entertaining children [e.g. Q129: "*Quarantine. Toddler. Smart speaker. Baby Shark.*"] and supporting children's socialization [e.g. Q148: "*Funnest part of my day is listening to my daughter argue with her Google Nest.*"]. For **Work**: managing distractions [e.g. Q137: "*Alexa reminders to keep me on task while I strategically slack off from work.*"], and simulating social environments [e.g. Q100: "*[...] piped some City Streets through my HomePod and WOW! it's kind of amazing at making my office feel less lonely.*"]. For **Housekeeping**: developing infrastructure for convenience [e.g. Q52: "*I've become more lazy during COVID. I have ordered an Amazon Echo and a load of smart plugs so I don't need to stand up to turn things on or off anymore.*"], and developing infrastructure for pleasure [e.g. Q67: "*There are smart lights in my living room now. Ready for a party but no one is invited until after lockdown...*"]. For **Leisure**: engaging in personal activities [Q143: "*I finally bought an Amazon Alexa during this pandemic just so I have someone to talk to.*"], and engaging in group activities [e.g. Q4: "*My husband asked Alexa how to do the Chicken Dance. She broke it down for him step by step. And we laughed like that was the funniest thing we've ever heard.*"]. Finally, for **Health**: tending to mental health [e.g. Q137: "*I've developed almost a new appreciation for smart home tech and automation under quarantine. All this money I thought I was wasting really did wonders for my mental health.*"], remote care-giving [e.g. Q181: "*Having Google Nest camera and Amazon Echo Show and the Harmony Hub have all made keeping in touch with my mom easier while her independent living apartment is on lockdown with no visitors.*"], and health 'spying' [e.g. Q12: "*I think Google Home is tracking if I have COVID-19. I smoke, so I occasionally will cough. Each time I cough my Google Home Mini lights up as if it were recording something.*"].

3 MOVING FORWARD

The goal of the presented initial research was to determine a set of clusters for IPA skills based on emerging contexts of use. We focused on data created during COVID-19 stay-at-home orders, since lockdown periods were marked by a unique convergence of all areas of life into a single environment. Through a qualitative analysis, we identified five 'area of life' themes in our data, as well as sub-themes of 'contexts of use' for each area of life, demonstrating how contexts of IPA-use differ between them. Moving along into the next stages of this ongoing work, we plan to investigate if the 'area of life'-based clustering of functionalities, observed in the initial study, would naturally manifest in the way users 'intuitively' cluster IPA skills.

REFERENCES

[1] Tawfiq Ammari, Jofish Kaye, Janice Y Tsai, and Frank Bentley. 2019. Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Trans. Comput. Hum. Interact.* 26, 3 (2019), 17–1.

[2] Sungeun An, Robert Moore, Eric Young Liu, and Guang-Jie Ren. 2021. Recipient Design for Conversational Agents: Tailoring Agent's Utterance to User's Knowledge. In *CUI 2021-3rd Conference on Conversational User Interfaces*. 1–5.

[3] John M Carroll and Mary Beth Rosson. 1987. Paradox of the active user. In *Interfacing thought: Cognitive aspects of human-computer interaction*. 80–111.

[4] Ramanathan Guha, Vineet Gupta, Vivek Raghunathan, and Ramakrishnan Srikant. 2015. User modeling for a personal assistant. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. 275–284.

[5] Ido Guy. 2016. Searching by talking: Analysis of voice queries on mobile web search. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 35–44.

[6] Philipp Kirschthaler, Martin Porcheron, and Joel E Fischer. 2020. What can I say? effects of discoverability in VUIs on task performance and user experience. In *Proceedings of the 2nd Conference on Conversational User Interfaces*. 1–9.

[7] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–31.

[8] Pierrick Milhorat, Stephan Schlögl, Gérard Chollet, Jerome Boudy, Anna Esposito, and Gianni Pelosi. 2014. Building the next generation of personal digital assistants. In *2014 1st international conference on advanced technologies for signal and image processing (atsip)*. IEEE, 458–463.

[9] Chelsea Myers, Anushay Furqan, Jessica Nebolsky, Karina Caro, and Jichen Zhu. 2018. Patterns for how users overcome obstacles in voice user interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–7.

[10] Edison Research and Triton Digital. 2021. The Infinite Dial 2021. http://www.edisonresearch.com/wp-content/uploads/2021/03/The-Infinite-Dial-2021.pdf

[11] Aafaq Sabir, Evan Lafontaine, and Anupam Das. 2022. Hey Alexa, Who Am I Talking to?: Analyzing Users' Perception and Awareness Regarding Third-party Alexa Skills. In *CHI Conference on Human Factors in Computing Systems*. 1–15.

[12] Shahin Sefati, Parsa Saadatpanah, Hassan Sayyadi, and Jan Neumann. 2018. Conversational content discovery via comcast X1 voice interface. In *Proceedings of the 12th ACM Conference on Recommender Systems*. 489–489.

[13] Nathan L Tenhundfeld, Hannah M Barr, HO Emily, and Kristin Weger. 2021. Is My Siri the Same as Your Siri? An Exploration of Users' Mental Model of Virtual Personal Assistants, Implications for Trust. *IEEE Transactions on Human-Machine Systems* (2021).

[14] Ryen W White. 2018. Skill discovery in virtual assistants. *Commun. ACM* 61, 11 (2018), 106–113.